

RESEARCH ARTICLE

WBC-AMNet: Automatic classification of WBC images using deep feature fusion network based on focalized attention mechanism

Ziyi Wang¹, Jiewen Xiao², Jingwen Li¹, Hongjun Li^{1*}, Luman Wang^{3*}

1 College of Science, Beijing Forestry University, Beijing, China, **2** College of Environmental Science and Engineering, Beijing Forestry University, Beijing, China, **3** Department of Health Informatics and Management, Peking University Health Science Center, Beijing, China

* lihongjun69@bjfu.edu.cn (HL); wangluman@bjmu.edu.cn (LW)



OPEN ACCESS

Citation: Wang Z, Xiao J, Li J, Li H, Wang L (2022) WBC-AMNet: Automatic classification of WBC images using deep feature fusion network based on focalized attention mechanism. PLoS ONE 17(1): e0261848. <https://doi.org/10.1371/journal.pone.0261848>

Editor: Mudassar Raza, COMSATS University Islamabad, Wah Campus, PAKISTAN

Received: August 18, 2021

Accepted: December 11, 2021

Published: January 27, 2022

Copyright: © 2022 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Our data are available at <https://github.com/wzy0730/minimal-dataset>.

Funding: This work is supported partly by the National Training Programs of Innovation and Entrepreneurship for Undergraduates (No. X202010022047), partly by the Fundamental Research Funds for the Central Universities, grant number No.2019SG05. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

The recognition and classification of White Blood Cell (WBC) play a remarkable role in blood-related diseases (i.e., leukemia, infections) diagnosis. For the highly similar morphology of different WBC subtypes, it is too confused to classify the WBC effectively and accurately for visual observation of blood cell smears. This paper proposes a Deep Convolutional Neural Network (DCNN) with feature fusion strategies, named WBC-AMNet, for automatically classifying WBC subtypes based on focalized attention mechanism. To obtain more localized attention of CNN, the fusion features of the first and the last convolutional layer are extracted by focalized attention mechanism combining Squeeze-and-Excitation (SE) and Gather-Excite (GE) modules. The new method performs successfully in classifying monocytes, neutrophils, lymphocytes, and eosinophils on the complex background with an overall accuracy of 95.66%, better than that of general CNNs. The multi-classification accuracy of WBC-AMNet with the background segmentation is over 98% in all cases. In addition, Gradient-weighted Class Activation Mapping (Grad-CAM) is employed to visualize the attention heatmaps of different feature maps.

Introduction

The analysis of White Blood Cell (WBC) images can assist clinical medicine experts in diagnosing many blood-related disorders such as leukopenia, Acute Leukemia (AL), agranulocytosis, etc. Importantly, AL is a malignant clonal disease of hematopoietic stem cells. Without special therapy, the average survival period is about three months, and even some patients died within a few days of diagnosis. AL is commonly classified into Acute Lymphoblastic Leukemia (ALL) and Acute Myelogenous Leukemia (AML) [1]. The survival rate of AML within five years is 40% [2], and in five cases in Europe, the annual survival rate of the disease is only 19% [3]. Therefore, the automated detection and classification of WBC sample images are of considerable reference value for leukemia diagnosis.

However, the presence of dyeing impurities and cytoplasm with low image contrast makes the microscopical differences between WBC more challenging to distinguish [4, 5].

Competing interests: The authors have declared that no competing interests exist.

In recent years, machine learning methods have been used for image classification of blood cells and have achieved excellent results in medicine. Standard high-performance classification methods and algorithms include Neural Network (NN), K-Nearest Neighbors (K-NN), Support Vector Machine (SVM), etc. The flexible neural tree algorithm of multi-classification NN cancer had an average accuracy rate of 98.6% on the mixed lineage leukemia [6]. The tumor diagnosis method based on the concept of biomarker association network could correctly classify all 72 samples for the WBC dataset [7]. Dichotomous classification of acute WBC samples achieved accuracy up to 86% [8]. Based on the multi-class SVM, an efficient hierarchical blood cell image recognition and classification method were proposed, with an average recall of 95.3% for six classifications of blood cells [9]. The advent of deep learning has led to experimentation with Convolutional Neural Network (CNN) in models for WBC classification [10].

At present, CNNs combining various methods have been successfully applied to the WBC classification [11, 12]. A recognition system, WBCsNet, based on deep convolution, was proposed to classify five categories on three different public WBC datasets with an accuracy of 96.1% [13]. A classification scheme involving CNN was proposed to classify 17092 images of normal peripheral blood cells with the best overall classification accuracy of 96.2% [14]. A CNN model with 32 feature maps achieved the accuracy of 88.25% and 81.74% in leukemia versus healthy and multi-classification of all subtypes, respectively [15]. Fifteen classifications were performed on 18365 peripheral blood smears using ResNeXt with an accuracy of more than 90% [16]. As powerful tools to assist physicians in diagnosing blood-related diseases, CNN algorithms still need further research on the generalizable properties and the explicit mechanisms of models detecting WBC images of blood smears.

Attention mechanism, as a standard feature extraction method, was widely used in deep learning and image classification in recent years due to its excellent performance and can be divided into Channel Attention Mechanism (CAM) and Spatial Attention Mechanism (SAM) [17]. We propose a method for automatically classifying WBC subtypes images in this paper. The new method, named WBC-AMNet, is based on attention mechanism and Deep CNN (DCNN), which well decreases the attention dispersing phenomenon due to the complex background of images.

The rest of this paper is presented below. Section 2 describes the material of WBC images in different contexts and elaborates the WBC-AMNet model in detail. The experimental results with different backgrounds and visual analysis are explored in section 3. A conclusion is drawn in the last section.

Materials and methods

WBC image datasets

In our experiment, the first WBC image dataset is from the Blood Cell Count Dataset (BCCD). BCCD contains the WBC images of eosinophils, monocytes, lymphocytes, and neutrophils cells. All images have complex background. Before our experiment, the enhanced 12515 images from BCCD have been divided into the training set, test set, and validation set. The details of the BCCD are described in Table 1.

Another public WBC images dataset (WBCs dataset) from the Kaggle repository incorporates 4358 raw WBC images segmented in single cells, which eliminates the interference of complex background. The WBCs dataset has been labeled as seven different WBC subtypes, and four of them are used in our experiment (Table 1). Images from the WBCs dataset are in RGB color space, 112×112 pixels, and JPEG format. During the training of our model, the dataset is divided into 60% for training, 20% for validation, and 20% for testing.

Table 1. Descriptions of BCCD and WBCs dataset.

Dataset	Description	Category	Division	Subtypes	Number
BCCD	12515 images 320 × 240	4	training set (9957)	neutrophils	2499
				monocytes	2478
				lymphocytes	2483
				eosinophils	2497
			test set(2487)	neutrophils	624
				monocytes	620
				lymphocytes	620
				eosinophils	623
			validation set(71)	neutrophils	48
				monocytes	4
				lymphocytes	6
				eosinophils	13
WBCs Dataset	4358 raw images 112 × 112	4	-	neutrophils	2025
				monocytes	576
				lymphocytes	1586
				eosinophils	171

<https://doi.org/10.1371/journal.pone.0261848.t001>

Several schematical images of two datasets with or without background segmentation are shown in Fig 1.

Methods

The pipeline of our approach for classifying the WBC images are described following. In the first, the WBC images of the online source are taken from blood smears under microscopes and labeled by experts.

And then, the images are pre-processed: all images are resized to 224 × 224 pixels to fit the model; the random rotation (by an angle of -14 ~ 15 degree), cropping, and flipping are used to eliminate the effects of irrelevant information and noise of images; and color distorting is conducted to make images clear. By color distorting, the brightness, contrast, saturation, and chromaticity of the image are adjusted by random factors taking values in (0, 1).

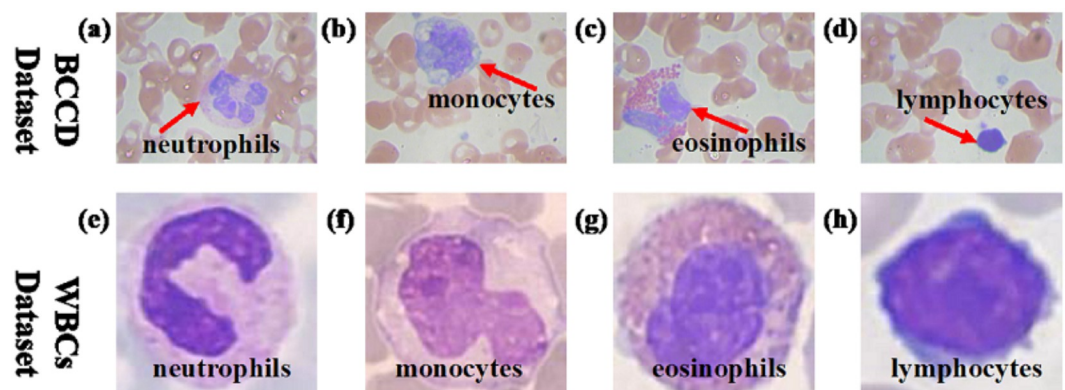


Fig 1. Sample images from BCCD (The first row) and the WBCs dataset (The second row). Among them, (a) and (e) are neutrophils, (b) and (f) are monocytes, (c) and (g) are eosinophils, and (d) and (h) are lymphocytes.

<https://doi.org/10.1371/journal.pone.0261848.g001>

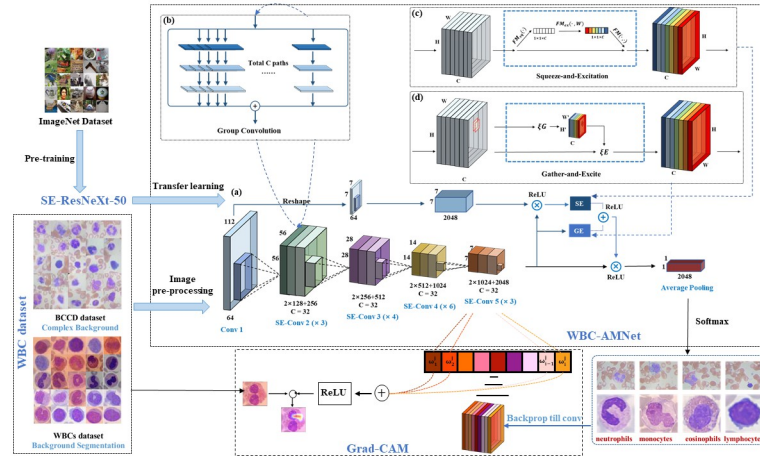


Fig 2. Flowchart of our method.

<https://doi.org/10.1371/journal.pone.0261848.g002>

After that, the images are input to the proposed WBC-AMNet with a part of parameters pre-trained on the ImageNet dataset to train and fine-tune the model. The WBC-AMNet is implemented the proposed focalized attention mechanism, and Grad-CAM is conducted to visualize the attention.

Finally, the classification results of WBC images are obtained and assist in diagnosing. Fig 2 depicts the flowchart of our method for the classification of different WBC subtypes.

At the training stage, the modes are performed on Baidu AI Studio platform with Tesla V100 GPU. The DCNNs model are implemented in the PaddlePaddle 1.7.2 deep learning framework of Python 3.7.

WBC-AMNet model

Our WBC-AMNet model is based on DCNN architecture with focalized attention mechanism. DCNN architecture mainly refers the WBCsNet [13] and uses the group convolution strategy.

Group convolution strategy. The group convolution strategy was employed in ResNeXt, while adopting the idea of VGG stacking and split-transform-merge of Inception [18]. Group convolution improves the accuracy while reducing the number of hyperparameters, sketch map as Fig 2b. Group convolution can not only significantly reduce the amount of model calculations [19], but also improve the accuracy of WBC-AMNet.

The focalized attention mechanism is mainly realized by Squeeze-and-Excitation (SE) module and Gather-Excite (GE) module.

Squeeze-and-Excitation (SE) module is a computational module constructed based on the CAM [20]. The core idea of the SE module is to model the interdependencies between feature channels, namely, use the global average pooling to squeeze the WBC feature map, perform a nonlinear transformation by excitation, finally superimpose on the input features, and recalibrate the feature channels by adaptive learning. The structure of the SE module is shown in Fig 2c.

Let $U = [u^1, u^2, \dots, u^C] \in R^{H \times W \times C}$ be the input, and $K = [k_1, k_2, \dots, k_c]$ denote the kernel set of filters in learning, k_i denotes the parameter of the i -th filter ($k_c = [k_c^1, k_c^2, \dots, k_c^C]$). After a series of transformations, we get $F = [f_1, f_2, \dots, f_c]$, then (Eq 1) [21]:

$$f_c = k_c * U = \sum_{s=1}^C k_c^s u^s \tag{1}$$

where $*$ denotes convolution. Thus, we can obtain the result F after SE module (Eq 2):

$$F(U) = \sum_{s=1}^C k_c^s \times u^s = \sum_{s=1}^C FM_{scale} ((FM_{ex}(FM_{sq}(u^s), W), u^s) \tag{2}$$

Based on the SE module, the Gather-Excite (GE) further exploits the feature context in the CNN by introducing a pair of operators ξG (step-wise deep convolution) and ξE [22]. The core of the GE module is to use different filters layer by layer on the feature map, which makes WBC-AMNet aggregate the features extracted by WBC accordingly. Gather can effectively aggregate feature responses on a large spatial scale, and then excite is used to redistribute the aggregated information to local features. The structure of the GE module is displayed in Fig 2d. After processing by GE module, WBC-AMNet focuses on local features more precisely and improves the feature extraction ability greatly.

Let V be the output ($V \in R^{[\frac{H}{\alpha}] \times [\frac{W}{\alpha}] \times C}$, $v \in R^{[\frac{H}{\alpha}] \times [\frac{W}{\alpha}]}$), then, the result G after GE module can be obtained as (Eq 3):

$$G = \sum_{s=1}^C \xi E(u^s, \xi G(u^s)) = \sum_{s=1}^C \xi E(u^s, \xi G(u^s \odot T_{\tau(v,\alpha)}^s)) = \sum_{s=1}^C u^s \odot g(\xi G(u^s \odot T_{\tau(v,\alpha)}^s)) \tag{3}$$

where \odot is the hadamard product, $\tau(v, \alpha) = \{\alpha v + \theta: \theta \in [-\lfloor \alpha - 1 \rfloor, \lfloor 2\alpha - 1 \rfloor^2/4]\}$, α : the selected range ratio. $T_{\{\cdot\}}$ denotes the tensor and g is the well-defined mapping.

The fused features of the first and last convolutional layers in the model are input to the SE module, while the feature maps of the last convolutional layer are input to the GE module. Then, the features of the SE module and GE module are fused to obtain the attentional features, and the attentional features are finally fused with the original features of the last convolutional layer. The output of the model is depended on the type of WBC in the dataset. The method of fine-tuning is implemented to obtain the optimal parameters of WBC-AMNet by transfer learning and gradient learning rate strategy.

Fig 2a depicts the focalized attention mechanism with DCNN architecture for WBC image classification, where the ReLU activation function is $ReLU(x) = \max(0, \omega^T x + b)$. \otimes is an element-wise multiplication operator, namely, the input X and the input Y are multiplied element-by-element, and the output elements at each position are stored in the returned result, $Out = X \otimes Y$. And \oplus is an element-wise add operator, namely, the input X and the input Y are added element-by-element, and the output elements at each position are saved in the returned result, $Out = X \oplus Y$.

The idea of focalized attention mechanism guides the WBC-AMNet construction base on the backbone of SE-ResNeXt implemented GE module and group convolution strategy (Table 2).

SE-ResNeXt. This module is used directly with residual networks of ResNeXt model to build SE-ResNeXt [20]. The innovation of the attention mechanism significantly improves the performance of the ResNeXt model with no additional calculative cost.

Attention visualization

WBC-AMNet only outputs numerical results such as accuracy, but it is difficult to intuitively understand the essential features and locations that the model finally extracts. To explain the effect of focalized attention mechanism in more vivid detail, we visualize the feature extraction and attention heatmap of WBC-AMNet using the Grad-CAM method [23]. Assume that the penultimate layer produces m features maps $F^m (F^m \in R^{H \times W}$ for any C) and F_{ij}^m is the activation

Table 2. Comparison of CNN structure between WBC-AMNet and other models.

stage	Output	ResNet-50 (32 × 4d)	SE-ResNeXt-50	WBC-AMNet	
conv1	112 × 112	7 × 7, 64, stride2	7 × 7, 64, stride2	7 × 7, 64, stride2	7 × 7, 2048, stride2
conv2	56 × 56	3 × 3, max pool, stride2	3 × 3, max pool, stride2	3 × 3, max pool, stride2	
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 256 \\ \text{fc}, [16, 256] \\ C = 32 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 256 \\ \text{fc}, [16, 256] \\ C = 32 \end{bmatrix} \times 3$	-
conv3	28 × 28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 512 \\ \text{fc}, [32, 512] \\ C = 32 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 512 \\ \text{fc}, [32, 512] \\ C = 32 \end{bmatrix} \times 4$	-
conv4	14 × 14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 1024 \\ \text{fc}, [64, 1024] \\ C = 32 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 1024 \\ \text{fc}, [64, 1024] \\ C = 32 \end{bmatrix} \times 6$	-
conv5	7 × 7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 1024 \\ 3 \times 3, 1024 \\ 1 \times 1, 2048 \\ \text{fc}, [128, 2048] \\ C = 32 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 1024 \\ 3 \times 3, 1024 \\ 1 \times 1, 2048 \\ \text{fc}, [128, 2048] \\ C = 32 \end{bmatrix} \times 3$	fc(SE(cov 1 + cov 5) + GE), [128, 2048]
				fc((SE(cov 1 + cov 5) + GE) + SE), [128, 2048]	
	1 × 1	global average pool 1000-d fc, softmax	global average pool 1000-d fc, softmax	global average pool 1000-d fc, softmax	

<https://doi.org/10.1371/journal.pone.0261848.t002>

of F^m at location (i, j). Grad-CAM obtains the gradient information of the score g^c for class C and uses the average value of all gradients as the weight of the feature map. After weighting the extracted features, the ReLU operation finally highlights the crucial regions in the WBC images through the class-discriminative localization map Grad-CAM $L^c_{Grad-CAM}$ (Eq 4). Grad-CAM does not require retraining of the proposed model, and it visualizes the local position in the WBC image which allows WBC-AMNet make the final decision.

$$L^c_{Grad-CAM} = \text{ReLU} \left(\sum_m \left(\frac{1}{M} \sum_i \sum_j \frac{\partial g^c}{\partial F^m_{ij}} \right) F^m \right) \tag{4}$$

Evaluation

This paper analyzes the performance of classification using indexes including Accuracy (Eq 5), Specificity (Eq 6), Precision (Eq 7), and F1-score (Eq 8). They are calculated as follows:

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{5}$$

$$\text{Specificity} = (\text{TN}) / (\text{TN} + \text{FP}) \tag{6}$$

$$\text{Precision} = (\text{TP}) / (\text{TP} + \text{FP}) \tag{7}$$

$$\text{F1-score} = (2\text{TP}) / (2\text{TP} + \text{FN} + \text{FP}) \quad (8)$$

Among them, True Positive (TP) indicates accurately identified positive data labels; False Positive (FP) indicates incorrectly identified positive data labels; True Negative (TN) indicates correctly identified negative data labels; False Negative (FN) indicates incorrectly identified negative data labels.

Experiment

Tri-classification results of BCCD

We perform a tri-classification analysis of WBC images from BCCD. Due to the presence of combined immune receptors based on T cell receptors in addition to T cells in neutrophils and monocytes, and they are derived from granulocyte monocyte progenitor cell [24, 35]. So we take monocytes and neutrophils cells as a set named MTD. The results of the search for epoch, batch size parameters, and the corresponding evaluation indexes are given in Table 3. Detailed data for WBC subtypes with different parameters are presented in S1 Table in S1 File. Fig 3 shows the change of objective function value in the training processing. As the number of iterations increases, the training accuracy improves rapidly in the initial stage, and then it converges to 1.00 gradually.

Our model reaches optimal performance when epoch = 20 and batch size = 32, at which point the accuracy reached 95.66%. Under the optimal parameters, we analyze in detail the recognition and classification ability of WBC-AMNet for three WBC subtypes. In Table 4, the

Table 3. Training results of tri-classification of BCCD images under different epoch and batch size.

Epoch	Batch size	WBC subtypes	Accuracy (%)	Specificity (%)	Precision (%)	F1-score (%)
15	32	3	90.71	90.71	90.74	90.70
20	16	3	94.93	94.94	95.03	94.93
20	32	3	95.66	94.70	95.67	95.66
25	32	3	95.13	95.13	95.14	95.11
30	32	3	93.81	93.81	93.97	93.69

<https://doi.org/10.1371/journal.pone.0261848.t003>

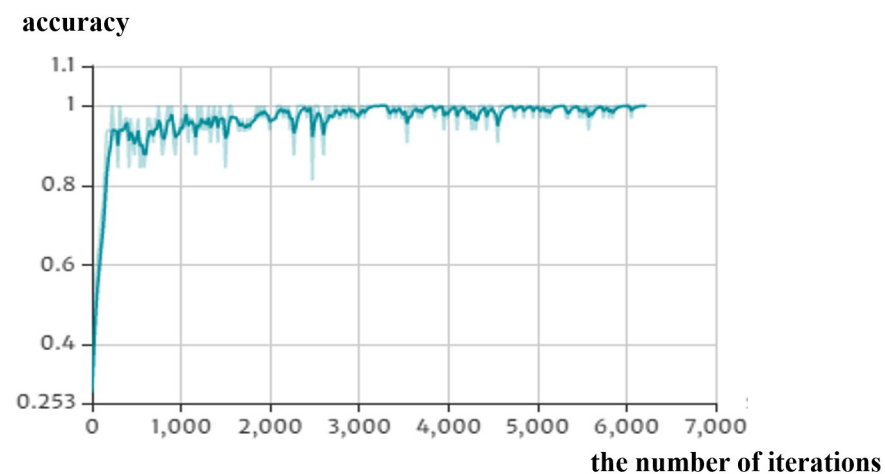


Fig 3. Classification accuracy versus the number of iterations in the training phase. (epoch = 20 and batch size = 32).

<https://doi.org/10.1371/journal.pone.0261848.g003>

Table 4. Training results when epoch = 20 and batch size = 32.

WBC subtypes	Accuracy (%)	Specificity (%)	Precision (%)	F1-score (%)
lymphocyte	100.00	95.50	100.00	100.00
MTD	95.50	91.65	95.81	95.65
eosinophil	91.65	100.00	91.07	91.36
total	95.66	94.70	95.67	95.66

<https://doi.org/10.1371/journal.pone.0261848.t004>

classification accuracy of lymphocytes is particularly outstanding, the precision of lymphocytes even reached 100%.

We use ROC curves and confusion matrices to visually show the classification performance of WBC-AMNet for each WBC subtype. The abscissa of the Receiver Operating Characteristic (ROC) curve is the FP Rate (FPR) and the ordinate is the TP Rate (TPR). The Area Under Curve (AUC) is defined as the area enclosed by the coordinate axis under the ROC curve. The larger the value of AUC, the better the performance of the model.

In Fig 4a, the solid lines of different colors represent the ROC curves of different WBC subtypes, and the blue dashed line represents the overall macroscopic average ROC curves. The AUC of the WBC-AMNet model is 0.99. All ROC curves demonstrate a high TPR and a low FPR. The ROC curves of each type are close to the upper left corner and far from the pure opportunity line. They indicate the solid tri-classification ability of the proposed model. The misclassification problem between MTD and eosinophils is reflected in the confusion matrix in Fig 4b. Both types of cells are easily confused with each other, whereas lymphocytes do not appear to be misclassified at all. The problem of misidentification between these two subtypes of WBC is also a common challenge for existing classifiers [11, 25].

We compare the tri-classification results obtained using our method with 10 general CNN models: VGG [26], ShuffleNetV2 [27], DPN [18], InceptionV4 [19], AlexNet [28], DistResNet [29], MobileNet-V1 [30], MobileNet-V2 [31], ResNet [29], SE-ResNeXt [32], as shown in Table 5. Detailed data for some of the models on the three WBC subtypes can be found in S2 Table in S1 File.

The accuracy of MobileNet-V1, MobileNet-V2, ResNet, DistResNet, and SE-ResNeXt are all over 93%, but the accuracy of WBC-AMNet is still significantly improved. The accuracy of WBC-AMNet is nearly two times higher than that of VGG, and other evaluation metrics also have significant differences. The simple structure of VGG makes it less practical for WBC

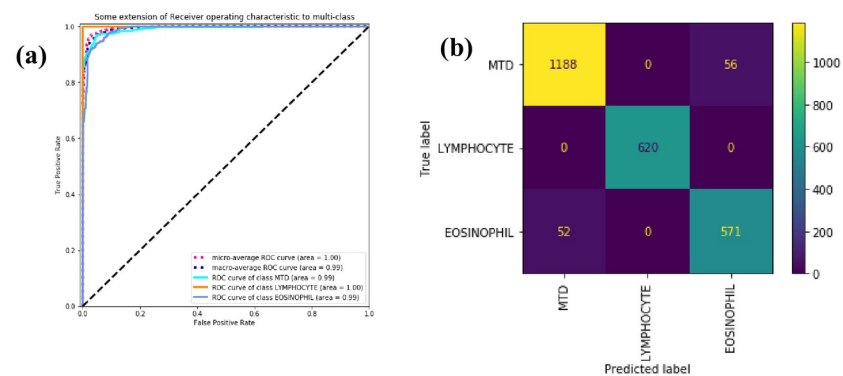


Fig 4. ROC curve and confusion matrix. (a) ROC curve of three subtypes of WBC. (b) Confusion matrix of three subtypes of WBC.

<https://doi.org/10.1371/journal.pone.0261848.g004>

Table 5. Training results when epoch = 20 and batch size = 32.

ID	CNN model	Accuracy (%)	Specificity (%)	Precision (%)	F1-score (%)
1	VGG	50.02	33.33	16.67	22.23
2	ShuffleNetV2	79.41	83.33	81.81	80.99
3	DPN	87.45	90.73	87.82	88.15
4	InceptionV4	90.59	88.25	93.48	90.27
5	AlexNet	93.00	91.73	93.85	92.55
6	DistResNet	94.29	92.80	95.74	93.93
7	MobileNet-V1	94.17	93.60	94.63	94.07
8	MobileNet-V2	94.45	94.45	94.41	94.37
9	ResNet	93.12	93.13	93.32	93.17
10	SE-ResNeXt	93.93	93.93	94.07	93.97
11	WBC-AMNet	95.66	94.70	95.67	95.66

<https://doi.org/10.1371/journal.pone.0261848.t005>

classification in complex background. Compared with ResNet and SE-ResNeXt, the accuracy of the proposed model is improved by about 2%, which is a surprising result. The high accuracy and other statistics demonstrate that introducing the feature fusion strategy and the focalized attention mechanism in complex background can significantly and effectively improve the WBC-AMNet classification ability. To visually analyze the training results of other models, we compare the confusion matrices in Fig 5.

VGG (Fig 5a) is a conventional CNN model, which identifies all WBC subtypes as MTD. Compared with MobilNetV2 (Fig 5b), WBC-AMNet improves the problem of misclassifying eosinophils as MTD. It reduces the number of eosinophils misclassified by nearly one third.

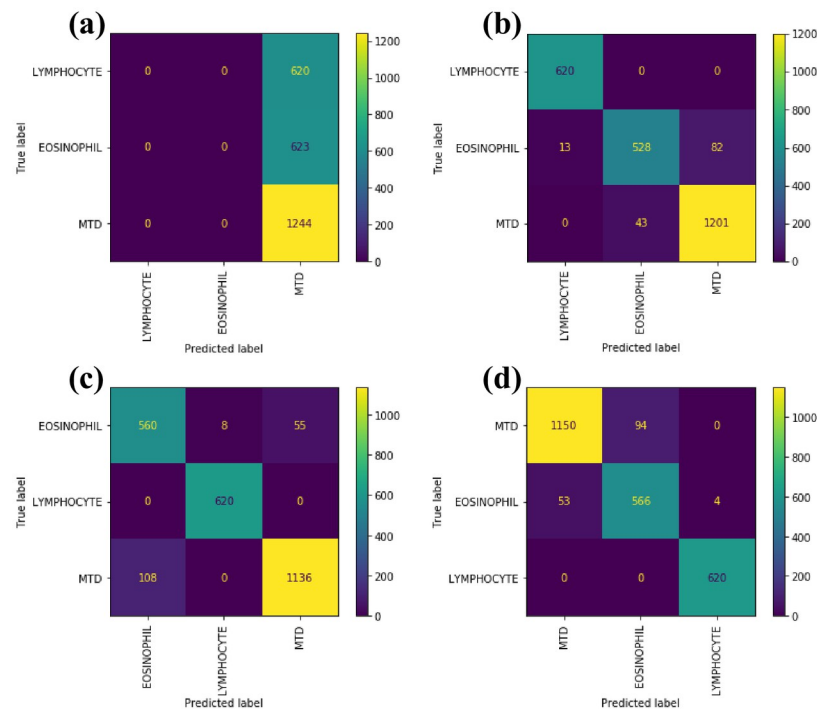


Fig 5. Confusion matrices of other CNN models. (a)VGG. (b)MobileNetV2. (c)ResNet. (d)SE-ResNeXt.

<https://doi.org/10.1371/journal.pone.0261848.g005>

ResNet (Fig 5c) addresses the problem of misclassifying MTD as eosinophils to a certain extent. The introduction of the focalized attention mechanism allowed WBC-AMNet to target attention to valuable features. SE-ResNeXt (Fig 5d) improves the problem of misclassifying MTD as eosinophils and shows unexpected results in predicting MTD. The combination of focalized attention mechanism and feature fusion allows WBC-AMNet to obtain local attention moreover.

Quad-classification results of BCCD

Monocytes and eosinophils are essential references for diagnosing diseases such as monocytic leukemia and an underlying allergic state, respectively [33, 34]. Accordingly, quad-classification is performed using WBC-AMNet for eosinophils, monocytes, lymphocytes, and neutrophils, with approximately 2480 training images and 620 test images for each subtype of WBC. Based on the results of the tri-classification parameter search, we refer to its optimal parameters (epoch = 20, batch size = 32), and the statistical results of different WBC subtypes are shown in Table 6. Due to a small proportion of monocytes and eosinophils are misclassified, resulting in their slightly lower accuracy. Detailed data for WBC subtypes with different parameters are presented in S3 Table in S1 File.

Comparing the results in Table 4, we find that lymphocytes still maintain a high classification accuracy. However, after reclassifying the MTD, the accuracy of eosinophils decreased, and the accuracy of neutrophils and monocytes is also not high. We speculate that a misclassification problem occurred [35]. Although the accuracy of neutrophils is high, the predicted result is not as well as it should have, resulting in a low F1-score. Conversely, although the accuracy of monocytes is low, it is incredibly predictive. The above phenomenon indicates that WBC-AMNet has a good classification ability for neutrophils, but the precision is higher for monocytes.

Fig 6a shows that the classification results are not satisfactory except for lymphocytes, reflected in the confusion matrix in Fig 6b. The quad-classification method identified monocytes and eosinophils as neutrophils several times, verifying our speculation. Other identification of WBCs as neutrophils is more numerous, but there are no cases of neutrophils identified as monocytes. Compared with monocytes, WBC-AMNet extracts the features of neutrophils more accurately.

Based on the 11 CNN models, the results and statistical data are shown in Table 7. Compared to Table 5, the classification ability of VGG is significantly improved on the quad-classification problems. The new model with the introduction of focalized attention mechanism has a significant improvement in accuracy compared to ResNet. The operation of feature fusion makes the classification accuracy of WBC-AMNet better than that of the best model SE-ResNeXt nowadays. Accuracy and other data intuitively reflect the important guiding significance of feature fusion for the model to extract features and process them.

Table 6. Training results of different WBC subtypes.

WBC subtypes	Accuracy (%)	Specificity (%)	Precision (%)	F1-score (%)
eosinophils	82.50	82.50	91.46	86.75
neutrophils	93.43	93.43	73.70	82.40
monocytes	84.03	84.03	98.67	90.77
lymphocytes	96.94	96.94	99.17	98.04
total	89.22	89.22	90.72	89.48

<https://doi.org/10.1371/journal.pone.0261848.t006>

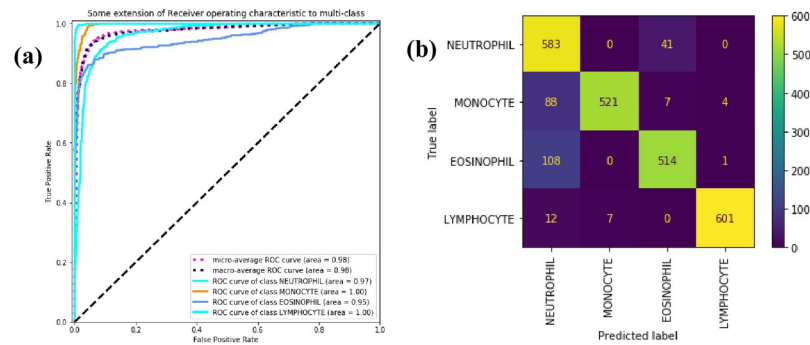


Fig 6. ROC curve and confusion matrix. (a) ROC curve of four subtypes of WBC. (b) Confusion matrix of four subtypes of WBC.

<https://doi.org/10.1371/journal.pone.0261848.g006>

Classification results of WBCs dataset

We use our method to classify the WBC images from WBCs dataset. These WBC images are all without complex background. In this section, we compare our method with 3 representative methods: From Tables 5 and 7, it can be found that MobileNetV2 has a higher accuracy rate. Comparing ResNet and SE-ResNeXt with WBC-AMNet, respectively, we will get the effect of introducing attention mechanism and GE module. On the premise of the same parameters (epoch = 20, batch size = 32), we choose MobileNetV2, ResNet and SE-ResNeXt to train and compare them with WBC-AMNet.

Tri-classification of WBCs dataset. First, the tri-classification results of different WBC subtypes are analyzed in Table 8. Slightly different from the BCCD, WBC-AMNet has a higher classification accuracy for neutrophils in the WBCs dataset, regardless of the model. Except for the intermediate cell, the classification accuracy of WBC-AMNet is above 99%, which is a satisfactory result. However, the accuracy of the intermediate cell is slightly lower, and we will reclassify it in the next section to further explore the reason. As shown in Fig 7, our model has the best performance among 4 methods in the view of 4 indexes: Accuracy, Specificity, Precision and F1-score.

The accuracy of WBC-AMNet combined with the focalized attention mechanism is nearly 4% higher than that of ResNet. WBC-AMNet also achieves more than 1% higher accuracy than SE-ResNeXt, not only due to the introduction of GE module but also thanks to the

Table 7. Statistical results of nine classic CNN models.

ID	CNN model	Accuracy (%)	Specificity (%)	Precision (%)	F1-score (%)
1	AlexNet	82.31	82.32	86.12	82.70
2	ShuffleNetV2	83.43	83.43	85.80	83.53
3	DPN	84.84	84.84	87.50	85.28
4	VGG	86.81	86.81	88.49	87.03
5	DistResNet	87.74	87.73	89.95	88.02
6	InceptionV4	87.94	87.94	90.41	88.23
7	MobileNet-V1	86.13	86.13	88.96	86.48
8	MobileNet-V2	88.82	88.82	90.85	89.02
9	ResNet	86.65	86.65	88.08	86.81
10	SE-ResNeXt	87.78	87.78	89.43	87.91
11	WBC-AMNet	89.22	89.22	90.72	89.47

<https://doi.org/10.1371/journal.pone.0261848.t007>

Table 8. Tri-classification results of images from WBCs dataset.

ID	CNN model	WBC subtypes	Accuracy(%)	Specificity(%)	Precision(%)	F1-score (%)
1	MobileNetV2	intermediate cell	84.46	84.46	89.29	86.81
		lymphocyte	94.75	94.65	96.78	95.71
		neutrophil	99.26	99.26	95.70	97.45
		total	95.06	95.06	95.00	95.00
2	ResNet	intermediate cell	77.70	77.70	95.04	85.50
		lymphocyte	95.91	95.91	96.83	96.37
		neutrophil	99.51	99.50	92.63	95.94
		total	94.48	94.48	94.58	94.32
3	SE-ResNeXt	intermediate cell	85.14	85.14	97.67	90.97
		lymphocyte	98.74	98.74	95.15	96.91
		neutrophil	99.75	99.75	98.05	98.90
		total	96.90	96.90	96.93	96.82
4	WBC-AMNet	intermediate cell	93.24	93.24	97.18	95.17
		lymphocyte	99.37	99.37	97.23	98.29
		neutrophil	99.01	99.01	99.26	99.13
		total	98.16	98.16	98.16	98.15

<https://doi.org/10.1371/journal.pone.0261848.t008>

operation of feature fusion. Regardless of the model, the classification accuracy of intermediate cells is lower than that of other cells. But for lymphocytes, the accuracy of the proposed model has increased, which is the main reason for the increase in total accuracy.

The AUCs of lymphocytes and neutrophils in Fig 8 are both 1.00, and the AUC of intermediate cells and the whole is 0.99. From Fig 8b, we can observe that a tiny number of intermediate cells (i.e. MTD) are still misclassified as lymphocytes, which is different from the conclusion of the BCCD. We suspect believed to be caused by different problems in different contexts misclassification. The confusion matrices of the four models depict in S5 Fig in S1 File.

In Fig 9, the solid lines of different colors represent the ROC curves of different WBC subtypes, and the blue dashed line represents the overall macroscopic average ROC curves. The AUC of MobileNetV2 (Fig 9a) is 0.98, the AUC of ResNet and SE-ResNeXt is 0.99. WBC-AMNet improves the classification of all three WBC subtypes. From Fig 9a and 9b, the light blue curve is slightly lower, which intuitively shows that the classification capabilities of

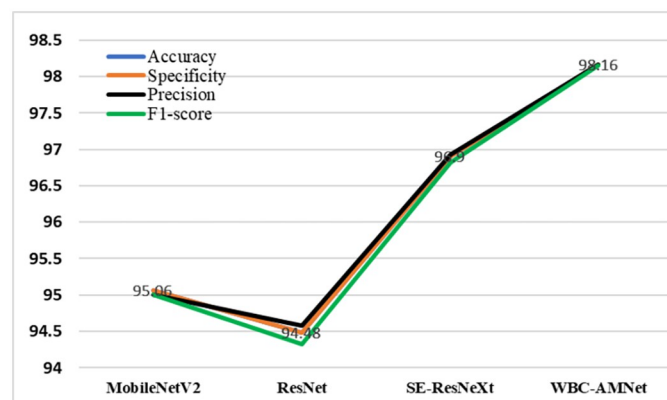


Fig 7. Tri-classification line chart of WBCs dataset.

<https://doi.org/10.1371/journal.pone.0261848.g007>

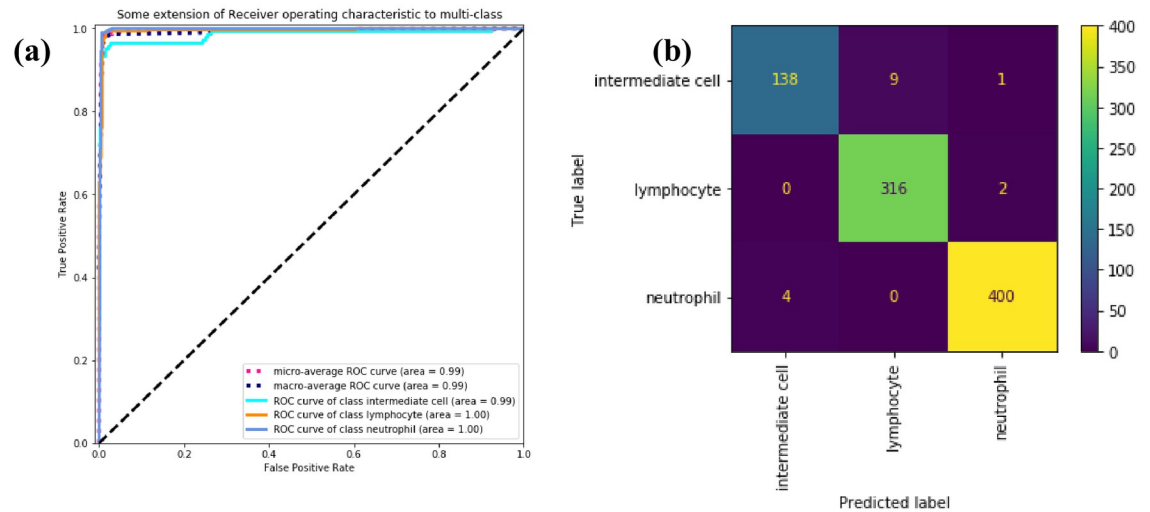


Fig 8. ROC curve and confusion matrix. (a) ROC curve of three subtypes of WBC. (b) confusion matrix of three subtypes of WBC.

<https://doi.org/10.1371/journal.pone.0261848.g008>

MobileNetV2 and ResNet for intermediate cells are little poor. Comparing Figs 8a to 9, SE-ResNeXt has improved over the previous two models, and WBC-AMNet is more stable for the classification of intermediate cells.

Quad-classification of WBCs dataset. The classification results and statistics of our method and three compared methods for the four WBC subtypes are listed in Table 9. The classification rate of WBC-AMNet in eosinophils becomes almost twice as high as that of MobileNetV2 and ResNet. Since both triple and quadruple classifications are improved, we conclude that WBC-AMNet gets the best performance among four methods.

MobileNetV2 has serious misclassification problems when recognizing eosinophils and monocytes, resulting in a low accuracy of these two subtypes of WBC. Except for eosinophil, ResNet has a higher accuracy for other WBC subtypes. However, since the number of images of eosinophils in the WBCs dataset is small, it has little effect on the overall accuracy. With the SE module, the accuracy of SE-ResNeXt has been significantly improved. Especially for eosinophils, SE-ResNeXt is about 35% higher than ResNet. Such a large increase in accuracy verifies the importance and effectiveness of using the attention mechanism strategy. The classification

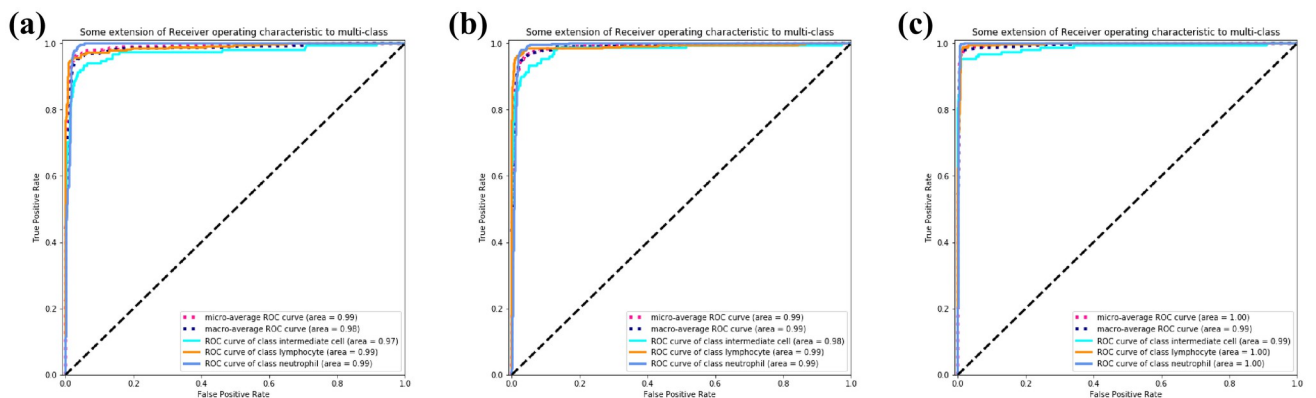


Fig 9. ROC curve. (a) MobileNetV2. (b) ResNet. (c) SE-ResNeXt.

<https://doi.org/10.1371/journal.pone.0261848.g009>

Table 9. Quad-classification results of images from WBCs dataset.

ID	CNN model	WBC subtypes	Accuracy(%)	Specificity(%)	Precision(%)	F1-score (%)
1	MobileNetV2	eosinophils	50.00	50.00	77.27	60.71
		neutrophils	98.02	98.02	95.66	96.83
		monocytes	69.82	69.83	84.38	76.42
		lymphocytes	97.79	97.78	91.42	94.50
		total	92.31	92.30	91.90	91.86
2	ResNet	eosinophils	61.76	61.76	91.30	73.68
		neutrophils	98.77	98.77	95.47	97.09
		monocytes	84.48	84.48	87.50	85.96
		lymphocytes	96.20	96.20	95.90	96.05
		total	94.49	94.49	94.40	94.32
3	SE-ResNeXt	eosinophils	97.06	97.06	97.06	97.06
		neutrophils	99.75	99.75	98.78	99.26
		monocytes	85.34	85.34	92.52	88.79
		lymphocytes	97.79	97.78	96.26	97.02
		total	97.02	97.01	96.96	96.97
4	WBC-AMNet	eosinophils	97.06	97.06	97.06	97.06
		neutrophils	99.75	99.75	99.02	99.38
		monocytes	95.69	95.69	97.37	96.52
		lymphocytes	97.78	97.78	98.10	97.94
		total	98.39	98.39	98.39	98.39

<https://doi.org/10.1371/journal.pone.0261848.t009>

accuracy of monocytes is still not satisfactory. By integrating the GE module, the recognition accuracy of WBC-AMNet in monocyte is improved by nearly 10% compared with SE-ResNeXt. Moreover, WBC-AMNet has an accuracy rate of over 95% for each WBC subtype. So far, we can come to the conclusion: WBC-AMNet has achieved effective WBC classification on WBCs dataset. Fig 10 shows that for quad-classification, our model, comparing to other 3 methods, still gets the higher scores of Accuracy, Specificity, Precision and F1-score.

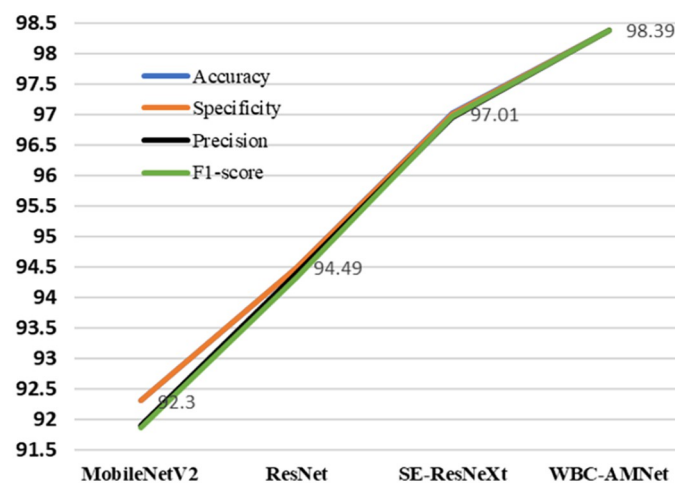


Fig 10. Quad-classification line chart of WBCs dataset.

<https://doi.org/10.1371/journal.pone.0261848.g010>

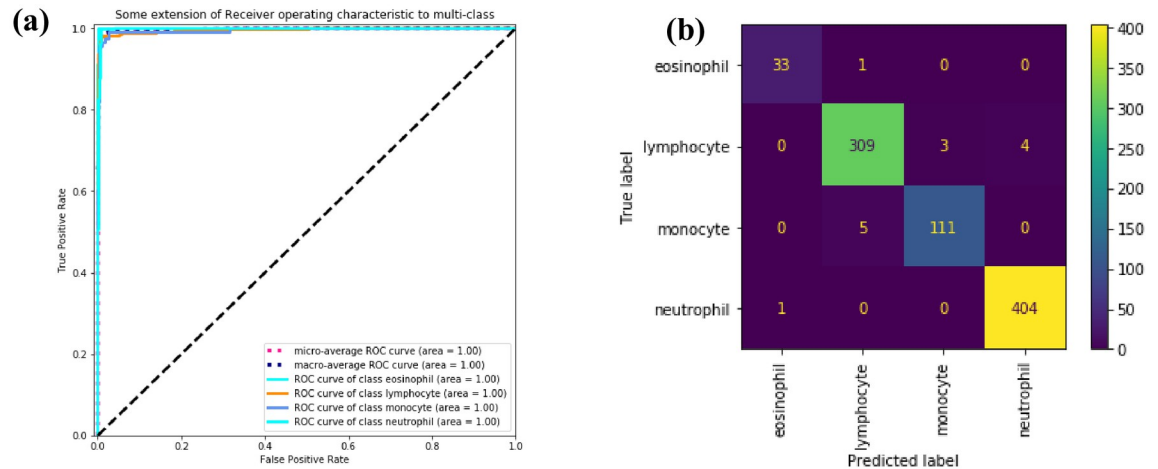


Fig 11. ROC curve and confusion matrix. (a) ROC curve of four subtypes of WBC. (b) Confusion matrix of four subtypes of WBC.

<https://doi.org/10.1371/journal.pone.0261848.g011>

In Fig 11a, the AUC value of 1.00 is reached for all WBC subtypes. There is almost no misclassification problem in Fig 11b. The confusion matrices of the four models are demonstrated in S6 Fig in S1 File.

In Fig 12, MobileNetV2 has an AUC of 0.98, ResNet has an AUC of 0.99, SE-ResNeXt and WBC-AMNet has an AUC of 1.00. The closer the AUC is to 1.00, the better the performance of the model. It can be seen from AUC that the performance of the model is improved after introducing the attention mechanism. Comparing Fig 12 with Fig 11a, the AUCs of MobileNetV2 and ResNet are lower on eosinophils and monocytes, the ROC of SE-ResNeXt on monocytes is slightly lower, and WBC-AMNet has reached 1.00 on all WBC subtypes, which means that WBC-AMNet is significantly improved compared to other CNN models.

Visualization analysis

The attention of different feature maps of WBC-AMNet is visualized in the background of single-cell segmentation. Firstly, the heatmap is obtained by a regular convolution operation. The first heatmap has highlighted regions spread over almost the whole WBC image and is very distracting. In order to focus the attention, the strategies of focalized attention mechanism and feature fusion are further introduced. The first and last convolutional layers in the model are

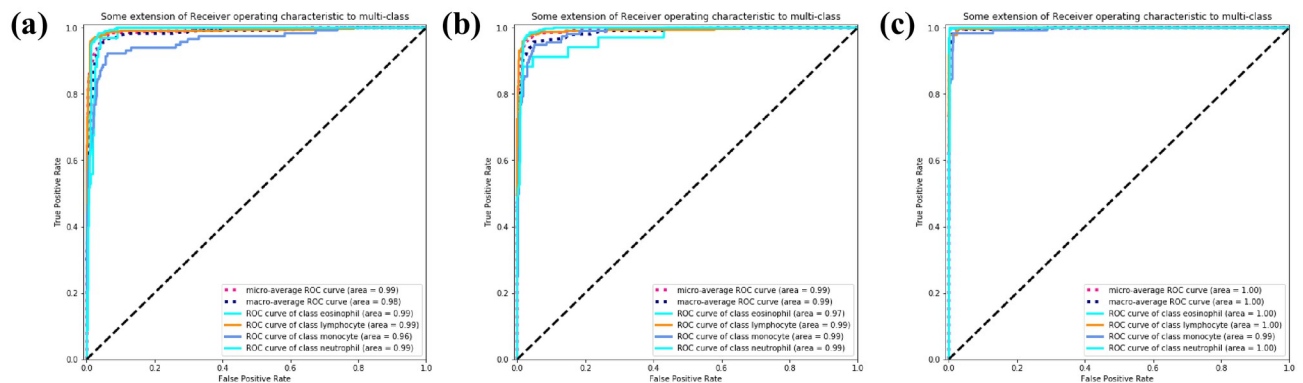


Fig 12. ROC curve. (a) MobileNetV2. (b) ResNet. (c) SE-ResNeXt.

<https://doi.org/10.1371/journal.pone.0261848.g012>

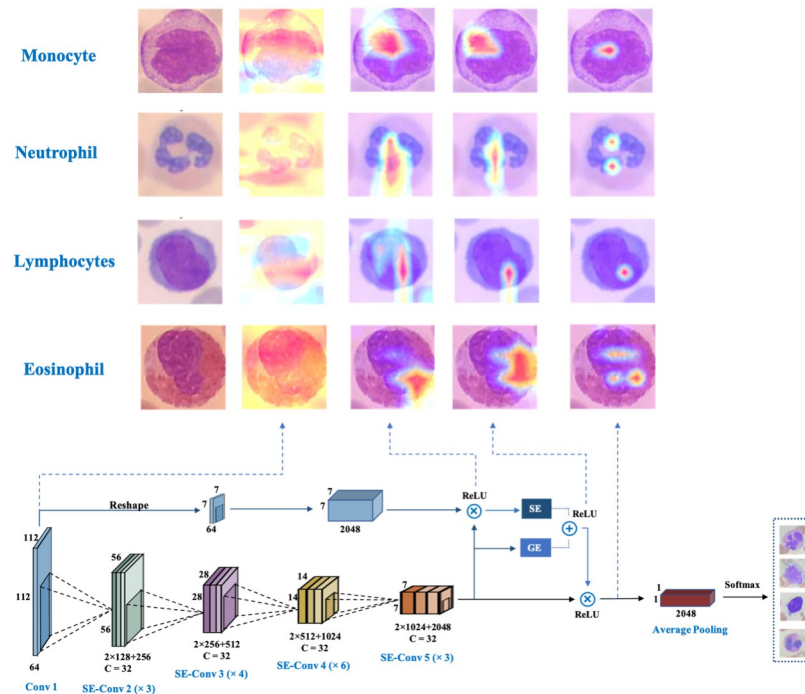


Fig 13. WBC-AMNet visualization analysis of attention to different feature maps.

<https://doi.org/10.1371/journal.pone.0261848.g013>

feature fused and fed into the SE module. At this point, the area of the highlighted region in the heatmap is significantly reduced, and the red part starts to accumulate in the cell nucleus. Then, the feature map of the last convolutional layer is input to the GE module and fused with the features of the SE and GE modules to obtain the attentional features. The attentional aggregation is slightly reduced, and almost all of them are on the WBC nuclei. Finally, the attention features are fused with the original features of the last convolutional layer. The final heatmap obtained reflects the superiority of the WBC-AMNet model. By implementing focalized attention mechanism and deep feature fusion, attention is highly focused on vital and partial locations of the WBC nuclei. Our proposed method extracts the effective critical information in the WBC cell nuclei and avoids the influence of too much redundant and invalid information on the results (Fig 13).

Conclusion

In this paper, we propose a new DCNN, WBC-AMNet, for automatic classification of WBC images based on focalized attention mechanism and deep feature fusion strategy. The attention of different feature maps of WBC-AMNet is visualized using the Grad-CAM method, which extracts the critical practical information from the WBC cell nuclei and avoids the influence of too much redundant and invalid information on the results. Experimental results show that WBC-AMNet gets the better performance than that of several existing models. Although the classification effect of our model is satisfactory, the mathematical mechanism of network architecture is still unclear. In the future, we intend to study the deep learning network from the perspective of mathematics and test more medical image data using our model.

Supporting information

S1 File. Supplementary material.
(PDF)

Acknowledgments

Baidu Ai Studio is acknowledged for AI training platform support.

Author Contributions

Conceptualization: Ziyi Wang, Jiewen Xiao.

Data curation: Ziyi Wang, Jingwen Li.

Formal analysis: Ziyi Wang.

Funding acquisition: Hongjun Li.

Investigation: Ziyi Wang.

Methodology: Ziyi Wang, Jiewen Xiao.

Project administration: Hongjun Li.

Resources: Ziyi Wang, Jiewen Xiao.

Software: Ziyi Wang, Jiewen Xiao.

Supervision: Hongjun Li, Luman Wang.

Validation: Ziyi Wang.

Visualization: Ziyi Wang, Jiewen Xiao.

Writing – original draft: Ziyi Wang, Jiewen Xiao.

Writing – review & editing: Ziyi Wang, Jiewen Xiao, Hongjun Li, Luman Wang.

References

1. Okikiolu J, Dillon R, Raj K. Acute leukaemia. *Medicine*. 2021; 49(5):274–281. <https://doi.org/10.1016/j.mpmed.2021.02.004>
2. Colvin GA, Elfenbein GJ. The latest treatment advances for acute myelogenous leukemia. *Medicine and health, Rhode Island*. 2003; 86(8):243–246. PMID: 14582219
3. Visser O, Trama A, Maynadié M, Stiller C, MarcosGragera R, Angelis RD, et al. Incidence, survival and prevalence of myeloid malignancies in Europe. *European Journal of Cancer*. 2012; 48(17):3257–3266. <https://doi.org/10.1016/j.ejca.2012.05.024> PMID: 22770878
4. Lu Y, Qin X, Fan H, Lai T, Li Z. WBC-Net: A white blood cell segmentation network based on UNet++ and ResNet. *Applied Soft Computing*. 2021; 101:107006. <https://doi.org/10.1016/j.asoc.2020.107006>
5. Zhao J, Zhang M, Zhou Z, Chu J, Cao F. Automatic detection and classification of leukocytes using convolutional neural networks. *Medical & Biological Engineering & Computing*. 2017; 55(8):1287–1301. <https://doi.org/10.1007/s11517-016-1590-x> PMID: 27822698
6. Lei X, Chen Y. Multiclass classification of microarray data samples with flexible neural tree. In: 2012 Spring Congress on Engineering and Technology; 2012. p. 1–4.
7. Wang H, Wong H, Zhu H, Yip T. A neural network-based biomarker association information extraction approach for cancer classification. *Journal of Biomedical Informatics*. 2009; 42(4):654–666. <https://doi.org/10.1016/j.jbi.2008.12.010> PMID: 19162234
8. Supardi N, Mashor M, Harun N, Bakri F, Hassan R. Classification of blasts in acute leukemia blood samples using k-nearest neighbour. In: 2012 IEEE 8th International Colloquium on Signal Processing and its Applications; 2012. p. 461–465.

9. Tai W, Hu R, Han C, Chen R, Tsai J. Blood cell image classification based on hierarchical SVM. In: 2011 IEEE International Symposium on Multimedia; 2011. p. 129–136.
10. Baydilli Y, Atila U. Classification of white blood cells using capsule networks. *Computerized Medical Imaging and Graphics*. 2020; 80:101699. <https://doi.org/10.1016/j.compmedimag.2020.101699> PMID: 32000087
11. Banik P, Saha R, Kim K. An automatic nucleus segmentation and CNN model based classification method of white blood cell. *Expert Systems with Applications*. 2020; 149:113211. <https://doi.org/10.1016/j.eswa.2020.113211>
12. Yang D, Zhao H, Han T, Kang Q, Ma J, Lu H. Leukocyte subtypes identification using bilinear self-attention convolutional neural network. *Measurement*. 2021; 173:108643. <https://doi.org/10.1016/j.measurement.2020.108643>
13. Shahin A, Guo Y, Amin K, Sharawi A. White blood cells identification system based on convolutional deep neural learning networks. *Computer Methods and Programs in Biomedicine*. 2019; 168:69–80. <https://doi.org/10.1016/j.cmpb.2017.11.015> PMID: 29173802
14. Acevedo A, Alf erez S, Merino A, Puigv ı L, Rodellar J. Recognition of peripheral blood cell images using convolutional neural networks. *Computer Methods and Programs in Biomedicine*. 2019; 180:105020. <https://doi.org/10.1016/j.cmpb.2019.105020> PMID: 31425939
15. Ahmed N, Yigit A, Isik Z, Alpkocak A. Identification of leukemia subtypes from microscopic images using convolutional neural network. *Diagnostics*. 2019; 9(3). <https://doi.org/10.3390/diagnostics9030104> PMID: 31450720
16. Matek C, Schwarz S, Spiekermann K, Marr C. Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks. *Nature Machine Intelligence*. 2019; 1(11):538–544. <https://doi.org/10.1038/s42256-019-0101-9>
17. Nan Y, Xi W. Classification of press plate image based on attention mechanism. In: 2019 2nd International Conference on Safety Produce Informatization (IICSPI); 2019. p. 129–132.
18. Chen Y, Li J, Xiao H, Jin X, Yan S, Feng J. Dual path networks. In: 31st Annual Conference on Neural Information Processing Systems (NIPS); 2017.
19. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-ResNet and the impact of residual connections on learning. In: 31st AAAI Conference on Artificial Intelligence; 2017. p. 4278–4284.
20. Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-Excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2020; 42(8):2011–2023. <https://doi.org/10.1109/TPAMI.2019.2913372> PMID: 31034408
21. Khan MM, ShorifUddin M, Parvez MZ, Nahar L. A squeeze and excitation ResNeXt-based deep learning model for Bangla handwritten compound character recognition. *Journal of King Saud University—Computer and Information Sciences*. 2021. <https://doi.org/10.1016/j.jksuci.2021.01.021>
22. Hu J, Shen L, Albanie S, Sun G, Vedaldi A. Gather-excite: Exploiting feature context in convolutional neural networks. In: 32nd Conference on Neural Information Processing Systems (NeurIPS); 2018.
23. Selvaraju R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision (ICCV); 2017. p. 618–626.
24. Attar A. Changes in the cell surface markers during normal hematopoiesis: a guide to cell isolation. *Global Journal of Hematology and Blood Transfusion*. 2014; 1(1):20–28. <https://doi.org/10.15379/2408-9877.2014.01.01.4>
25. Wang Q, Wang J, Zhou M, Li Q, Wen Y, Chu J. A 3D attention networks for classification of white blood cells from microscopy hyperspectral images. *Optics & Laser Technology*. 2021; 139:106931. <https://doi.org/10.1016/j.optlastec.2021.106931>
26. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations (ICLR); 2015.
27. Ma N, Zhang X, Zheng H, Sun J. ShuffleNet V2: Practical guidelines for efficient CNN architecture design. In: Proceedings of the European Conference on Computer Vision (ECCV); 2018.
28. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: NeurIPS; 2012.
29. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 770–778.
30. Howard A, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. In: arXiv preprint arXiv:1704.04861; 2017.

31. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. MobileNetV2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2018.
32. Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017. p. 5987–5995.
33. Khan A, Eker A, Chefranov A, Demirel H. White blood cell type identification using multi-layer convolutional features with an extreme-learning machine. *Biomedical Signal Processing and Control*. 2021; 69:102932. <https://doi.org/10.1016/j.bspc.2021.102932>
34. Baker F, Silverton R. Introduction to haematology. Introduction to Medical Laboratory Technology (Fifth Edition): Butterworth-Heinemann. 1976; p. 549–558.
35. Fuchs T, Puellmann K, Wang C, Han J, Beham AW, Neumaier Michael, et al. Trilineage sequencing reveals complex TCR & Transcriptomes in Neutrophils and Monocytes Alongside T Cells. *Genomics, Proteomics & Bioinformatics*. 2021. <https://doi.org/10.1016/j.gpb.2019.02.004>